

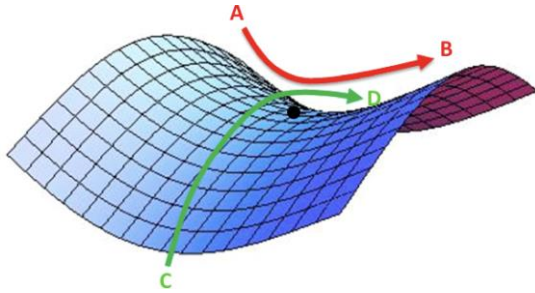
## Ch2 Berry Phase

### Gauss–Bonnet theorem and topology

In geometry, if a manifold  $M_1$  can be adiabatically deformed into  $M_2$ , we say that they have the same topology. Otherwise, we say that they are topologically different. To distinguish different manifolds, mathematicians developed an object, which is called an “index” (a topological index). It is an integer number. For objects with the same topology, the index takes the same value. Otherwise, the values are different. For 2D closed manifold, the index is the Euler characteristic:

$$\chi_M = \frac{1}{2\pi} \iint_M K dS$$

To define the curvature for a curve, we use a circle to fit the curve around one point on the curve. The inverse radius  $\kappa = 1/R$  gives us the curvature. For a manifold, one can draw lots of curves at one point and one can get the curvature for each of these curves. Among all these curvatures, the largest and smallest are known as the principle curvatures  $\kappa_1$  and  $\kappa_2$ . The Gaussian curvature is the product of they two. For a sphere,  $\kappa_1 = \kappa_2 = 1/R$ , so  $K = \kappa_1 \kappa_2 = 1/R^2$ . For a saddle point, the surface curves up along one direction and curves down along another direction, so that we have  $\kappa_1 > 0$  and  $\kappa_2 < 0$ . As a result,  $K = \kappa_1 \kappa_2 < 0$ .



For 2D closed manifold, the total Gaussian curvature is always  $2\pi$  times an integer. Therefore, the value of  $\chi_M$  is quantized to

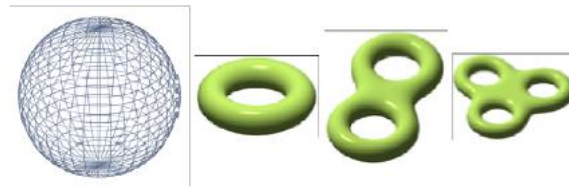


integer.

Sphere:  $\chi_M = 2$

Torus:  $\chi_M = 0$

Double torus:  $\chi_M = -2$



From left to right: sphere, torus, double torus and triple torus.

Triple torus:  $\chi_M = -4$

Again, it is important to emphasize here that this integral is a topological index, **only if we are considering a closed manifold which has no boundary**. Otherwise, it is not quantized and it is not a topological index.

## 1.1 Abelian Berry phase and Berry connection

We'll describe the Berry phase arising for a general Hamiltonian which we write as

$$H(x^a; \lambda^i)$$

The Hamiltonian depends on two different kinds of variables. The  $x^a$  are the degrees of freedom of the system. They are typically things like the *positions or spins* of particles. In contrast, the other variables  $\lambda^i$  are the parameters of the Hamiltonian, which can be fixed by external elements.

There is a theorem in quantum mechanics called the adiabatic theorem. This states that if we place a system in a non-degenerate energy eigenstate and vary parameters sufficiently slowly, then the system will cling to that energy eigenstate. We pick some values for the parameters  $\lambda$  and place the system in a specific energy eigenstate  $|\psi\rangle$  which, for simplicity, we will take to be the ground state. We assume this ground state is unique. Now we slowly vary the parameters  $\lambda$ . The Hamiltonian changes so, of course, the ground state also changes; it is  $|\psi(\lambda(t))\rangle$ .

The idea of the Berry phase arises in the following situation: we vary the parameters  $\lambda$  but, ultimately, we put them back to their starting values. This means that we trace out a closed path in the space of parameters.

If we started in the ground state, we also end up in the ground state. The only thing left uncertain is the phase of this new state

$$|\psi\rangle \rightarrow e^{i\gamma} |\psi\rangle$$

We often think of the overall phase of a wavefunction as being unphysical. But that's not the case here because this is a phase difference. So what is the phase  $e^{i\gamma}$ ? There are two contributions. The first is simply the dynamical phase  $e^{iEt/\hbar}$  that is there for any energy eigenstate, even if the parameters don't change. But there is also another, less obvious contribution to the phase. This is the Berry phase.

### 1.1.1 Adiabatic theorem and the Berry connection

Let us assume that we have a Hamiltonian  $H(\lambda(t))$  that depends on time dependent parameters  $\lambda(t)$ . Started from the adiabatic approximation, we now turn to the limit that the Hamiltonian changes very slowly. Here "slow" is again relative to the energy splitting, implying that these

parameters evolve slowly  $\hbar \frac{\partial \lambda(t)}{\partial t} \ll \Delta E$  where  $\Delta E$  is the minimal gap between the ground state and the first excited state at any given time  $t$ . To see this, we want to solve the time-dependent Schrödinger equation

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H(\lambda(t)) |\psi(t)\rangle$$

For simplicity, we will assume that  $H$  has a discrete spectrum.

$$H(n(\lambda)) |\psi\rangle = E_n(\lambda) |n(\lambda)\rangle$$

Let's place ourselves in one of these energy eigenstates. Now vary the parameters  $\lambda^i$ . If  $\lambda^i$  are changed suitably slowly, then the system will cling to the energy eigenstate  $|n(\lambda(t))\rangle$  that we started off in. We expand the solution to the Schrödinger equation in a basis of instantaneous energy eigenstates,

$$|\psi(t)\rangle = \sum_m a_m(t) e^{i\xi_m(t)} |m(\lambda(t))\rangle$$

Here  $a_m(t)$  are coefficients that we wish to determine, while  $\xi_m(t)$  is the usual energy-dependent phase factor

$$\xi_m(t) = -\frac{1}{\hbar} \int_0^t dt' E_m(t')$$

Substitute our ansatz above,  $|\psi(t)\rangle$ , into the Schrödinger equation to find

$$\sum_m \left[ \dot{a}_m e^{i\xi_m} |m(\lambda)\rangle + a_m e^{i\xi_m} \frac{\partial}{\partial \lambda^i} |m(\lambda)\rangle \dot{\lambda}^i \right] = 0$$

where we've cancelled the two terms which depend on  $E_n$ . Taking the inner product with  $\langle n(\lambda) |$  gives

$$\begin{aligned} \dot{a}_n &= - \sum_m a_m e^{i(\xi_m - \xi_n)} \langle n(\lambda) | \frac{\partial}{\partial \lambda^i} |m(\lambda)\rangle \dot{\lambda}^i \\ &= i a_n \mathcal{A}_i(\lambda) \dot{\lambda}^i - \sum_{m \neq n} a_m e^{i(\xi_m - \xi_n)} \langle n(\lambda) | \frac{\partial}{\partial \lambda^i} |m(\lambda)\rangle \dot{\lambda}^i \end{aligned}$$

In the second line, we've singled out the  $m = n$  term and defined

$$\mathcal{A}_i(\lambda) = -i \langle n | \frac{\partial}{\partial \lambda^i} |n\rangle$$

This is called the Berry connection.

First, we need to deal with the second term in  $\dot{a}_n$ . Differentiate the original definition of  $H(n(\lambda)) |\psi\rangle = E_n(\lambda) |n(\lambda)\rangle$  with respect to  $\lambda$ ,

$$\frac{\partial H}{\partial \lambda^i} |m\rangle + H \frac{\partial}{\partial \lambda^i} |m\rangle = \frac{\partial E_m}{\partial \lambda^i} |m\rangle + E_m \frac{\partial}{\partial \lambda^i} |m\rangle$$

Now take the inner product with  $\langle n|$  where  $n \neq m$  to find

$$(E_m - E_n) \langle n| \frac{\partial}{\partial \lambda^i} |m\rangle = \langle n| \frac{\partial H}{\partial \lambda^i} |m\rangle$$

This means that the second term in  $\dot{a}_n$  is proportional to

$$\langle n| \frac{\partial}{\partial \lambda^i} |m\rangle \dot{\lambda}^i = \langle n| \frac{\partial H}{\partial \lambda^i} |m\rangle \frac{\dot{\lambda}^i}{E_m - E_n}$$

The adiabatic theorem holds when the change of parameters  $\dot{\lambda}^i$  is much smaller than the splitting of energy levels  $E_m - E_n$ . In this limit, we can ignore this term. From  $\dot{a}_n$ , we're then left with

$$\dot{a}_n = i a_n \mathcal{A}_i \dot{\lambda}^i$$

So

$$a_n = C_n \exp \left( i \int_0^t dt' \mathcal{A}_i(\lambda(t')) \dot{\lambda}^i \right)$$

where  $C_n$  are constants.

This is the *adiabatic theorem*. If we start at time  $t = 0$  with  $a_m = \delta_{mn}$ , so the system is in a definite energy eigenstate  $|n\rangle$ , then the system remains in the state  $|n(\lambda)\rangle$  as we vary  $\lambda$ . This is true as long as  $\hbar \dot{\lambda}^i \ll \Delta E$ , so that we can drop the term  $\langle n| \frac{\partial}{\partial \lambda^i} |m\rangle \dot{\lambda}^i$ .

### 1.1.2 The Berry phase

We can see the picture clearer after doing some more calculation. If we start at  $t = 0$  in the ground state, we will always stay in the instantaneous ground state. However, along the way we will pick up a phase

$$e^{i\xi_m(t)} |\psi(\lambda)\rangle: i\hbar \frac{\partial}{\partial t} e^{i\xi_m(t)} |\psi(\lambda)\rangle = H(\lambda) e^{i\xi_m(t)} |\psi(\lambda)\rangle$$

Multiplying this expression from the left with  $\langle \psi(\lambda) | e^{-i\xi_m(t)}$ , we obtain

$$\frac{\partial \langle \psi(\lambda) |}{\partial t} = i \langle \psi(\lambda) | \frac{\partial}{\partial \lambda^i} | \psi(\lambda) \rangle \dot{\lambda}^i - \frac{1}{\hbar} E_m(\lambda)$$

Integrating this equation leads to

$$\xi(t) - \xi(0) = \int_{\lambda(0)}^{\lambda(t)} \langle \psi(\lambda) | \frac{\partial}{\partial \lambda^i} | \psi(\lambda) \rangle d\lambda - \frac{1}{\hbar} \int_0^t dt' E_0(t')$$

The first integral is the Berry phase while the second integral is the dynamical phase.

To highlight the distinction between these two contributions, suppose that we vary the parameters  $\lambda$  but, finally we put them back to their starting values. This means that we trace out a closed path  $C$  in the space of parameters. The second contribution can now be written as

$$e^{i\gamma} = \exp\left(i \oint_C \mathcal{A}_i(\lambda) d\lambda^i\right)$$

In contrast to the energy-dependent phase, this does not depend on the time taken to make the journey in parameter space. Instead, it depends only on the path we take through parameter space.

There is some very pretty geometry underlying the Berry phase. This is an example of a kind of object that you've seen before: it is like the gauge potential in electromagnetism!

In the relativistic form of electromagnetism, we have a gauge potential  $A_\mu(x)$  where  $\mu = 0, 1, 2, 3$  and  $x$  are coordinates over Minkowski spacetime. All physics remains invariant under the gauge transformation

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \omega$$

for any function  $\omega(x)$ . This is an example of a kind of object that you've seen before: it is like the gauge potential in electromagnetism! We should compute the field strength

$$F_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu}$$

This contains the electric and magnetic fields. It is invariant under gauge transformations.

Now let's compare this to the Berry connection  $\mathcal{A}_i(\lambda)$ . Of course, this no longer depends on the coordinates of Minkowski space; instead it depends on the parameters  $\lambda^i$ . The number of these parameters is arbitrary; let's suppose that we have  $d$  of them. This means that  $i = 1, \dots, d$ . We could pick a different phase for every choice of parameters  $\lambda$ ,

$$|n'(\lambda)\rangle = e^{i\omega(\lambda)} |n(\lambda)\rangle$$

for any function  $\omega(\lambda)$ . If we compute the Berry connection arising from this new choice, we have

$$\begin{aligned} \langle n'(\lambda) | &= e^{-i\omega(\lambda)} \langle n(\lambda) | \rightarrow \mathcal{A}'_i = -i \langle n' | \frac{\partial}{\partial \lambda^i} | n' \rangle = -ie^{-i\omega(\lambda)} \langle n(\lambda) | \frac{\partial}{\partial \lambda^i} [e^{i\omega(\lambda)} | n(\lambda) \rangle] \\ &= -ie^{-i\omega(\lambda)} \langle n(\lambda) | \left[ ie^{i\omega(\lambda)} \frac{\partial \omega}{\partial \lambda^i} | n(\lambda) \rangle + e^{i\omega(\lambda)} \frac{\partial}{\partial \lambda^i} | n(\lambda) \rangle \right] \\ &= \frac{\partial \omega}{\partial \lambda^i} - i \langle n(\lambda) | \frac{\partial}{\partial \lambda^i} | n(\lambda) \rangle \end{aligned}$$

$$\mathcal{A}'_i = -i\langle n' | \frac{\partial}{\partial \lambda^i} | n' \rangle = \mathcal{A}_i + \frac{\partial \omega}{\partial \lambda^i}$$

This takes the same form as the gauge transformation.

Following the analogy with electromagnetism, we might expect that the physical information in the Berry connection can be found in the gauge-invariant field strength which, mathematically, is known as the curvature of the connection,

$$\mathcal{F}_{ij}(\lambda) = \frac{\partial \mathcal{A}_i}{\partial \lambda^j} - \frac{\partial \mathcal{A}_j}{\partial \lambda^i}$$

It's certainly true that  $\mathcal{F}$  contains some physical information about our quantum system. In fact, it's possible to write the Berry phase in terms of the field strength using the higher-dimensional version of Stokes' theorem

$$e^{i\gamma} = \exp\left(-i \oint_C \mathcal{A}_i(\lambda) d\lambda^i\right) = \exp\left(-i \oint_S \mathcal{F}_{ij} dS^{ij}\right) \quad (0.1)$$

where  $S$  is a two-dimensional surface in the parameter space bounded by the path  $C$ .

### 1.1.3 Quantum anholonomy and electromagnetism

Just like the vector potential, the definition of Berry connection  $\mathcal{A}(\mathbf{k})$  depends on a particular choice of the person making the calculation. If you decide to multiply the quantum state by a phase,  $|\psi(\mathbf{k})\rangle \rightarrow \exp[i\lambda(\mathbf{k})]|\psi(\mathbf{k})\rangle$ , then you get that the Berry connection as  $\mathcal{A}(\mathbf{k}) \rightarrow \mathcal{A}(\mathbf{k}) + \nabla_k \lambda A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \omega$ . However, when you take the integral of  $\mathcal{A}(\mathbf{k})$  on a closed path, the result is independent of  $\lambda$ . That's why the Berry phase is only meaningful for closed paths. In the following table, we summarize the analogy to gauge field:

| Electromagnetism                          | Quantum anholonomy 异和乐 (微分几何概念)        |
|---|--|
| vector potential $\mathbf{A}(\mathbf{r})$ | Berry connection $\mathbf{A}(\lambda)$ |
| magnetic field $\mathbf{B}(\mathbf{r})$   | Berry curvature $\mathcal{F}(\lambda)$ |
| Magnetic (Dirac) monopole                 | degenerate point (Chern number)        |
| magnetic charge                           | Berry index                            |
| magnetic flux $\Phi(C)$                   | Berry phase $\gamma(C)$                |

## 1.2 Examples

### 1.2.1 A spin in a magnetic field

The standard example of the Berry phase is very simple. It is a spin, with a Hilbert space consisting of just two states. The spin is placed in a magnetic field  $\vec{B}$ , and we will assume that the magnetic field is uniform in space, so there is no coupling to the spatial degrees of freedom. The Hamiltonian is

$$H = -\vec{B} \cdot \vec{\sigma} + B$$

with  $\vec{\sigma}$  the triplet of Pauli matrices and  $B = |\vec{B}|$ . The offset ensures that the ground state always has vanishing energy. Indeed, this Hamiltonian has two eigenvalues: 0 and  $+2B$ . We denote the ground state as  $|\downarrow\rangle$  and the excited state as  $|\uparrow\rangle$ ,

$$H|\downarrow\rangle = 0, \quad H|\uparrow\rangle = 2B|\uparrow\rangle$$

Note that these two states are non-degenerate as long as  $\vec{B} \neq 0$ .

We are going to treat the magnetic field as the parameters, so that  $\lambda^i \equiv \vec{B}$  in this example. The specific form of  $|\uparrow\rangle$  and  $|\downarrow\rangle$  will depend on the orientation of  $\vec{B}$ . We write the magnetic field  $\vec{B}$  in spherical polar coordinates

$$\vec{B} = \begin{pmatrix} B \sin \theta \cos \phi \\ B \sin \theta \sin \phi \\ B \cos \theta \end{pmatrix}$$

with  $\theta \in [0, \pi]$  and  $\phi \in [0, 2\pi)$ . The Hamiltonian then reads

$$H = -B \begin{pmatrix} \cos \theta - 1 & e^{-i\phi} \sin \theta \\ e^{+i\phi} \sin \theta & -\cos \theta - 1 \end{pmatrix}$$

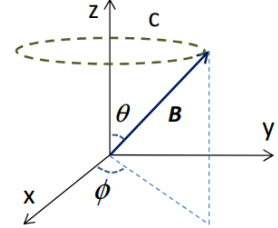
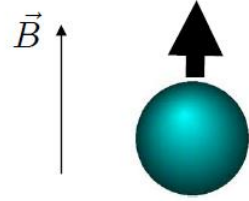
In these coordinates, two normalized eigenstates are given by

$$|\downarrow\rangle = \begin{pmatrix} e^{-i\phi} \sin \theta / 2 \\ -\cos \theta / 2 \end{pmatrix}, \quad |\uparrow\rangle = \begin{pmatrix} e^{-i\phi} \cos \theta / 2 \\ \sin \theta / 2 \end{pmatrix}$$

Both states describe the ground state. These states play the role of our reference state  $|n(\lambda)\rangle$  that we had in our general derivation. Note, however, that they are not well defined for all values of  $\vec{B}$ . When we have  $\theta = \pi$ , the angular coordinate  $\phi$  is not well defined. This means that  $|\downarrow\rangle$  and  $|\uparrow\rangle$  don't have well defined phases. For example,  $|\downarrow\rangle$  is singular at  $\theta = \pi$  (or at the south-pole), while  $|\uparrow\rangle$  is singular at the north-pole. This kind of behavior is typical of systems with non-trivial Berry phase.

We can easily compute the Berry phase arising from these states (staying away from  $\theta = \pi$  to be on the safe side). We have

$$\mathcal{A}_\theta = -i\langle\downarrow|\frac{\partial}{\partial\theta}|\downarrow\rangle = 0, \quad \mathcal{A}_\phi = -i\langle\downarrow|\frac{\partial}{\partial\phi}|\downarrow\rangle = -\sin^2(\theta/2)$$



The resulting Berry curvature in polar coordinates is

$$\mathcal{F}_{\theta\phi} = \frac{\partial \mathcal{A}_\phi}{\partial \theta} - \frac{\partial \mathcal{A}_\theta}{\partial \phi} = -\sin \theta$$

This is simpler if we translate it back to cartesian coordinates where the rotational symmetry is more manifest. It becomes

$$\mathcal{F}_{ij}(\vec{B}) = -\epsilon_{ijk} \frac{B^k}{2|\vec{B}|^3}$$

But this is interesting. It is a magnetic monopole! Of course, it's not a real magnetic monopole of electromagnetism: those are forbidden by the Maxwell equation. Instead it is, rather confusingly, a magnetic monopole in the space of magnetic fields.

We know from electromagnetism, the magnetic monopole has magnetic charge given by  $g = -1/2$ . This means that the integral of the Berry curvature over any two-sphere  $\mathbf{S}^2$  which surrounds the origin is

$$\int_{S^2} \mathcal{F}_{ij} dS^{ij} = 4\pi g = -2\pi$$

Let's now see what this has to do with magnetic monopoles. We place our particle, with electric charge  $q$ , in the background of a magnetic monopole with magnetic charge  $g$ . We keep the magnetic monopole fixed, and let the electric particle undergo some journey along a path  $C$ . We will ask only that the path  $C$  avoids the origin where the magnetic monopole is sitting. Upon returning, the particle picks up a phase  $e^{iq\alpha/\hbar}$  with

$$\alpha = \oint_C \mathbf{A} \cdot d\mathbf{x} = \int_S \mathbf{B} \cdot d\mathbf{S}$$

$S$  is the area enclosed by  $C$ . Using the fact that  $\int_{S^2} \mathbf{B} \cdot d\mathbf{S} = g$ , if the surface  $S$  makes a solid angle  $\Omega$ , this phase can be written as

$$\alpha = \frac{\Omega g}{4\pi}$$

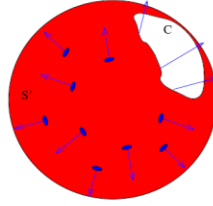
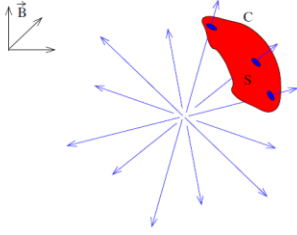
Instead of integrating over  $S$ , it is equally valid to calculate the phase by integrating over  $S'$ . The solid angle formed by  $S'$  is  $\Omega' = 4\pi - \Omega$ . The phase is then given by

$$\alpha' = -\frac{(4\pi - \Omega)g}{4\pi}$$

where the overall minus sign comes because the surface  $S'$  has the opposite orientation to  $S$ . So we must have  $e^{iq\alpha/\hbar} = e^{iq\alpha'/\hbar}$ .

$$\begin{aligned}\frac{q\alpha}{\hbar} &= \frac{q\alpha'}{\hbar} + 2\pi n \\ \frac{q(\alpha - \alpha')}{\hbar} &= q \left( \frac{\Omega g + 4\pi g - \Omega g}{4\pi} \right) / \hbar = \frac{qg}{\hbar} \\ e^{iq(\alpha - \alpha')\hbar} &= e^{iqg} \\ qg &= 2\pi n\hbar\end{aligned}$$

where  $n \in \mathbf{Z}$ . This is the famous Dirac quantization condition. The smallest such magnetic charge has  $n = 1$ , then  $g = 2\pi\hbar/q$ .



Suppose that the surface  $S$ , bounded by  $C$ , makes a solid angle  $\Omega$ . Then, we have

$$e^{i\gamma} = \exp\left(-i \int_S \mathcal{F}_{ij} dS^{ij}\right) = \exp\left(\frac{i\Omega}{2}\right)$$

Note, however, that there is an ambiguity in this computation. We could choose to form  $S$  as shown in the left hand figure. But we could equally well choose the surface  $S'$  to go around the back of the sphere, as shown in the right-hand figure. In this case, the solid angle formed by  $S'$  is  $\Omega' = 4\pi - \Omega$ . Computing the Berry phase using  $S'$  gives

$$e^{i\gamma'} = \exp\left(-i \int_{S'} \mathcal{F}_{ij} dS^{ij}\right) = \exp\left(\frac{-i(4\pi - \Omega)}{2}\right) = e^{i\gamma} = \exp\left(\frac{i\Omega}{2}\right) \quad (0.2)$$

where the difference in sign in the second equality comes because the surface now has opposite orientation. Note, however, that this agreement requires that the charge of the monopole is  $2g \in \mathbf{Z}$ . In the context of electromagnetism, this was Dirac's original argument for **the quantization of monopole charge**. This quantization extends to a general Berry curvature  $\mathcal{F}_{ij}$  with an arbitrary number of parameters: the integral of the curvature over any closed surface must be quantized in units of  $2\pi$ ,

$$\int \mathcal{F}_{ij} dS^{ij} = 2\pi C$$

The integer  $C \in \mathbf{Z}$  is called the **Chern number**.

## Singularity

For a closed surface enclosing a magnetic monopole, no matter what gauge one uses, the vector potential must have some singularities.

If  $A$  is a non-singular function on a closed manifold, the magnetic flux through this manifold must be zero. To prove this, we cut the manifold into two parts  $D_I$  and  $D_{II}$ .

The magnetic flux through  $D_I$  is  $\iint_{D_I} dS \cdot B = \iint_{D_I} dS \cdot \nabla \times A = \int_{\partial D_I} dl \cdot A$

Here we use the Stokes' theorem and  $\partial D_I$  is the edge of  $D_I$ .

The magnetic flux through  $D_{II}$  is  $\iint_{D_{II}} dS \cdot B = \iint_{D_{II}} dS \cdot \nabla \times A = \int_{\partial D_{II}} dl \cdot A$

Therefore, the total magnetic flux is  $\oint_M B \cdot dS = \iint_{D_I} dS \cdot B + \iint_{D_{II}} dS \cdot B = \int_{\partial D_I} dl \cdot A + \int_{\partial D_{II}} dl \cdot A$

Notice that the edge of  $D_I$  and  $D_{II}$  are the same curve but their directions are opposite  $\partial D_I = -\partial D_{II}$

$$\oint_M B \cdot dS = \int_{\partial D_I} dl \cdot A - \int_{\partial D_I} dl \cdot A = 0$$

The only way to have nonzero magnetic flux here is to have some singular vector potential. If  $A$  is singular (for all any gauge choice), we must use at least two different gauges to cover the whole manifold.

If we use  $A_I$  for  $D_I$  uses  $A_{II}$  for  $D_{II}$ , we get

$$\oint_M B \cdot dS = \int_{\partial D_I} dl \cdot A_I - \int_{\partial D_I} dl \cdot A_{II} = \int_{\partial D_I} dl \cdot (A_I - A_{II})$$

Here, I emphasize that this singularity in  $A$  is not a physical singularity. If you measure any physical observables, there is nothing singular here. The singularity lies in  $A$ , which is NOT a physical observable.

## The first Chern number

Later we will show that the same conclusion holds for the Berry curvature. For insulators, if the Berry connection can be defined for any  $k$  in the BZ without singularity. We can show that

$$\oint_{BZ} \mathcal{F} d\vec{k} = \int_{\partial D_I} dl \cdot A - \int_{\partial D_I} dl \cdot A = 0$$

To have nonzero Berry curvature, the Berry connection must have some singular point in the  $k$ -space. As a result, we will need to define different Bloch wave functions for different regions of the  $k$ -space and these Bloch waves are connected by a gauge transformation at the boundary.

$$|u_{n,k}^{II}\rangle = e^{i\phi_n(k)} |u_{n,k}^I\rangle$$

We know that the Berry connection changes under this phase shift, so the Berry connections in the different regions must be different also

$$\mathcal{A}_n^{II} = \mathcal{A}_n^I + \nabla_{\vec{k}} \phi_n(k)$$

This singularity again is not a physical singularity. Because we use different  $\mathcal{A}$  for different regions, we have

$$\oint_{BZ} \mathcal{F} d\vec{k} = \int_{\partial D_I} dl \cdot (\mathcal{A}_I - \mathcal{A}_{II})$$

which could be none zero.

The quantization of the flux is only true for a closed manifold (e.g. a sphere). For an open manifold with boundaries, the flux is in general unquantized (e.g. part of a sphere).

The same is true for the Berry curvature. Only if the integral is over the whole BZ, the total Berry curvature is quantized. In other words, the Berry curvature for insulator is quantized. But for metal, it is not.

### The Berry curvature and the Chern number

For an insulator, we can use the Bloch waves to define a curvature in the 2-dimensional k-space. For Bloch waves  $\psi_{n,\vec{k}}(\vec{r}) = u_{n,\vec{k}}(\vec{r}) \exp(i\vec{k} \cdot \vec{r})$ . We can define the Berry curvature for band  $n$  at momentum point  $\vec{k}$

$$\mathcal{F}_n(\vec{k}) = \iint_{unit\ cell} [\nabla_{\vec{k}} u_{n,\vec{k}}(\vec{r})]^* \times \nabla_{\vec{k}} u_{n,\vec{k}} d\vec{r} = \epsilon_{ij} \iint_{unit\ cell} [\partial_{k_i} u_{n,\vec{k}}(\vec{r})]^* \partial_{k_j} u_{n,\vec{k}}(\vec{r}) d\vec{r}$$

Here the integral is over a unit cell. The  $\times$  in the first line is the cross product (notice that the gradient  $\nabla_{\vec{k}}$  gives us a vector in the k-space).

For mathematicians, this Berry curvature is the same object as the Gaussian curvature, and thus the total Berry curvature shall also give us a topological index, which is indeed the case. This topological index is known as the Chern number:

$$C_n = \frac{1}{2\pi} \oint_{BZ} \mathcal{F}_n(\vec{k}) d\vec{k}$$

For each band, we can define such a topological index for this band.

For an insulator, we can define the topological index for the insulator as the total Chern number of all filled bands

$$C = \sum_{\text{filled bands}} C_n$$

If  $C = 0$ , we have a trivial insulator with no edge states ( $\sigma_{xx} = \sigma_{xy} = 0$ ). If  $C \neq 0$ , we call this insulator a Chern insulator.

For a metal, we need to sum over all fully occupied bands as well as the occupied states in a partially filled bands. Therefore, we need to sum up both the filled valence band and the partially filled conduction band. For an insulator we only have fully filled bands, so we only need to account for the filled valence band.

It is important to notice that for the Gaussian curvature, the total Gaussian curvature is quantized only if when we consider closed 2D manifolds without edges. For Berry curvature, the same is true. If the integral is over the whole BZ (no boundary), we have quantized Chern numbers. If the integral is over part of a BZ, then we will not get an integer. For a metal, the conduction band is only partially filled and the integral is over the filled states only (i.e. the Fermi sea). The Fermi sea has a boundary, which is the Fermi surface. If our manifold has a boundary (the Fermi surface), the integral is NOT quantized. This is the reason why metals don't have a quantized Hall conductivity, but insulators do.

### Dirac Quantization, Gauss–Bonnet theorem and the Chern number

From the mathematical point of the view, the following three objects are the same thing (same fiber bundles): the magnetic field  $B$ , the Berry curvature  $F$ , and the Gaussian curvature of  $K$  (geometry).

$\oint_M B \cdot dS = \oint_M B_n \cdot dS = \frac{\hbar}{2e} n$  quantized:  $n$  is an integer, known as the magnetic charge, which measures the number of magnetic monopole inside  $M$ .

$\oint_M K dS = 2\pi\chi_M$  quantized:  $\chi_M$  is an even integer, known as the Euler characteristic, which measures the topological nature of the manifold  $M$ .

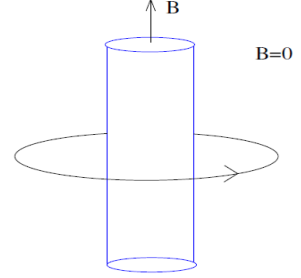
$\oint_{BZ} \mathcal{F} d\vec{k} = 2\pi C$  quantized:  $C$  is an integer, known as the TKNN invariant or the Chern number, which measures the quantized Hall conductivity for a topological insulator.

#### 1.2.2 Particles moving around a flux tube

In our course on electromagnetism, we learned that the gauge potential  $A_\mu$  is unphysical: the physical quantities that affect the motion of a particle are the electric and magnetic fields. This statement is certainly true classically. Quantum mechanically, it requires some caveats. This is the

subject of the Aharonov-Bohm effect. As we will show, aspects of the Aharonov-Bohm effect can be viewed as a special case of the Berry phase.

To see how the gauge potential  $\vec{A}$  can affect the physics, consider the set-up shown on the right. We have a solenoid of area  $A$ , carrying magnetic field  $\vec{B} = (0, 0, B)$  and therefore magnetic flux  $\Phi = BA$ . Outside the solenoid the magnetic field is zero. However, the vector potential is not. This follows from Stokes' theorem which tells us that the line integral outside the solenoid is given by



$$\oint \vec{A} \cdot d\vec{r} = \int \vec{B} \cdot d\vec{S} = \Phi$$

This is simply solved in cylindrical polar coordinates by  $A_\phi = \frac{\Phi}{2\pi r}$

Now consider a charged quantum particle restricted to lie in a ring of radius  $r$  outside the solenoid. The only dynamical degree of freedom is the angular coordinate  $\phi \in [0, 2\pi)$ . The Hamiltonian is

$$H = \frac{1}{2m} (p_\phi + eA_\phi)^2 = \frac{1}{2mr^2} \left( -i\hbar \frac{\partial}{\partial \phi} + \frac{e\Phi}{2\pi} \right)^2$$

We'd like to see how the presence of this solenoid affects the particle. The energy eigenstates are simply

$$\psi = \frac{1}{\sqrt{2\pi r}} e^{in\phi}, n \in \mathbf{Z}$$

where the requirement that is single valued around the circle means that we must take  $n \in \mathbf{Z}$ . Plugging this into the time independent Schrödinger equation  $H\psi = E\psi$

$$E = \frac{1}{2mr^2} \left( \hbar n + \frac{e\Phi}{2\pi} \right)^2 = \frac{\hbar^2}{2mr^2} \left( n + \frac{\Phi}{\Phi_0} \right)^2, n \in \mathbf{Z}$$

Note that if  $\Phi$  is an integer multiple of the quantum of flux  $\Phi_0 = 2\pi\hbar/q$ , then the spectrum is unaffected by the solenoid. But if the flux in the solenoid is not an integer multiple of  $\Phi_0$ , then the spectrum gets shifted. We see that the energy of the particle knows about the flux  $\Phi$  even though the particle never goes near the region with magnetic field.

Suppose now that we turn off the solenoid and place the particle in the  $n = 0$  ground state. Then we very slowly increase the flux. By the adiabatic theorem, the particle remains in the  $n = 0$  state. But, by the time we have reached  $\Phi = \Phi_0$ , it is no longer in the ground state. It is now in the state that we previously labelled  $n = 1$ . Similarly, each state  $n$  is shifted to the next state,  $n + 1$ . This is an example of a phenomenon is called spectral flow: *under a change of parameter—in this case*

$\Phi$  —the spectrum of the Hamiltonian changes, or “flows”. As we change increase the flux by one unit  $\Phi_0$  the spectrum returns to itself, but individual states have morphed into each other.

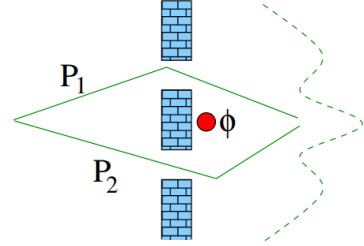
### 1.2.3 The Aharonov-Bohm Effect

Consider a particle that obeys the free Schrödinger equation,

$$\frac{1}{2m}(-i\hbar\nabla - q\mathbf{A})^2\psi = E\psi$$

We can formally remove the gauge field by writing

$$\psi(\mathbf{x}) = \exp\left(\frac{iq}{\hbar} \int^{\mathbf{x}} \mathbf{A}(\mathbf{x}') \cdot d\mathbf{x}'\right) \phi(\mathbf{x})$$



where the integral is over any path. Crucially, however, in the double-slit experiment there are two paths, P1 and P2. The phase picked up by the particle due to the gauge field differs depending on which path is taken. The phase difference is given by

$$\Delta\theta = \frac{q}{\hbar} \int_{P1} \mathbf{A} \cdot d\mathbf{x} - \frac{q}{\hbar} \int_{P2} \mathbf{A} \cdot d\mathbf{x} = \frac{q}{\hbar} \oint \mathbf{A} \cdot d\mathbf{x} = \frac{q}{\hbar} \int \mathbf{B} \cdot d\mathbf{S}$$

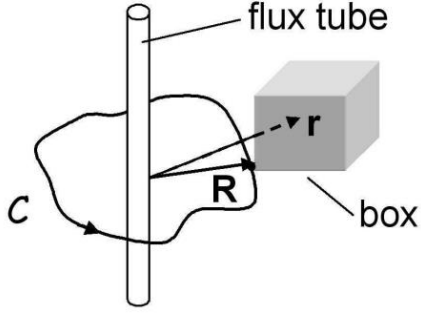
Note that neither the phase arising from path P1, nor the phase arising from path P2, is gauge invariant. However, the difference between the two phases is gauge invariant. This is the Aharonov-Bohm phase,  $\exp(iq\Phi/\hbar)$ , an extra contribution that arises when charged particles move around magnetic fields.

The Aharonov-Bohm phase manifests in the interference pattern seen on the screen. As  $\Phi$  is changed, the interference pattern shifts, an effect which has been experimentally observed. Only when  $\Phi$  is an integer multiple of  $\Phi_0$  is the particle unaware of the presence of the solenoid.

The situation described above smells like the Berry phase story. Consider a set-up like the solenoid where the magnetic field is localized to some region of space. A magnetic field is confined in a tube with flux  $\Phi$  and a box located at  $\mathbf{R}$  in which a particle of charge  $q$  are confined. However, this time we restrict the particle to lie in a small box. There can be some interesting physics going on inside the box; The fact that the box is “small” means that the gauge potential is approximately constant inside the box. If we place the center of the box at position  $\vec{x} = \vec{X}$ , then the Hamiltonian of the system is then

$$H(\vec{p}, \vec{x} - \vec{X}) = \frac{1}{2m} \left( -i\hbar\nabla + e\vec{A}(\vec{X}) \right)^2 + V(\vec{x} - \vec{X})$$

The corresponding wave functions, for a vanishing vector potential, are of the form  $\psi(\vec{x} - \vec{X})$  with the energy  $E$  independent of  $\vec{X}$ . We start by placing the center of the box at position  $\vec{x} = \vec{X}_0$  where we'll take the gauge potential to vanish:  $\vec{A}(\vec{X}_0) = 0$ . The energies are independent of the vector potential since we can always do a gauge transformation to ensure that  $\vec{A}$  vanishes at any point of our choosing



$$\psi(\vec{x} - \vec{X}_0)$$

which is localized around  $\vec{x} = \vec{X}_0$  as it should be. When the flux is non-zero, we can choose as basis states  $|n(\vec{X})\rangle$ , satisfying  $H(\vec{p} - q\vec{A}(\vec{x}), \vec{x} - \vec{X}) = E|n(\vec{X})\rangle$

Now we slowly move the box in some path in space. In doing so, the gauge potential  $\vec{A}(\vec{x} - \vec{X})$  experienced by the particle changes.

$$\begin{aligned} \langle r|n(\vec{X})\rangle &= \psi(\vec{x} - \vec{X}) \\ &= \exp\left(-\frac{ie}{\hbar} \int_{\vec{x}=\vec{X}_0}^{\vec{x}=\vec{X}} \vec{A}(\vec{x}) \cdot d\vec{x}\right) \psi(\vec{x} - \vec{X}_0) \end{aligned}$$

where the integral is performed along a path contained in the box. Then we take the box in a loop  $C$  and bring it back to where we started. The wavefunction comes back to

$$\psi(\vec{x} - \vec{X}_0) \rightarrow e^{i\gamma} \psi(\vec{x} - \vec{X}_0) \quad \text{with} \quad e^{i\gamma} = \exp\left(-\frac{ie}{\hbar} \oint_C \vec{A}(\vec{x}) \cdot d\vec{x}\right)$$

Comparing this to our general expression for the Berry phase, we see that in this particular context the Berry connection  $\vec{\mathcal{A}}(\vec{X})$  is actually identified with the electromagnetic vector potential,

$$\begin{aligned} \vec{\mathcal{A}}(\vec{X}) &= -Im[\langle n(\vec{X}) | \partial_{\vec{X}} n(\vec{X}) \rangle] \\ \vec{\mathcal{A}}(\vec{X}) &= -Im \iiint d^3\vec{x} \psi^*(\vec{x} - \vec{X}) \left[ -\frac{ie}{\hbar} \vec{A}(\vec{X}) \psi(\vec{x} - \vec{X}) + \partial_{\vec{X}} \psi(\vec{x} - \vec{X}) \right] \\ \vec{\mathcal{A}}(\vec{X}) &= \frac{e}{\hbar} \vec{A}(\vec{x} = \vec{X}) \end{aligned}$$

Therefore the Berry phase measures the flux of the magnetic field across the interior of the solenoid. In general, if a particle of charge  $q$  goes around a region containing flux  $\Phi$ , it will pick up an Aharonov-Bohm phase

$$\exp(iq\Phi/\hbar)$$

There is an experiment which exhibits the Berry phase in the Aharonov-Bohm effect. The wavefunction splits so the particle, in essence, travels through both. Except now, we hide a solenoid carrying magnetic flux  $\Phi$  behind the wall. The wavefunction of the particle is prohibited

from entering the region of the solenoid, so the particle never experiences the magnetic field  $\vec{B}$ . Nonetheless, as we have seen, the presence of the solenoid induces a phase difference  $e^{i\gamma}$  between particles that take the upper slit and those that take the lower slit. This phase difference manifests itself as *a shift to the interference pattern seen on the screen*. Note that when  $\Phi$  is an integer multiple of  $\Phi_0$ , the interference pattern remains unchanged; it is only sensitive to the fractional part of  $\Phi/\Phi_0$ .

In classical electromagnetism, the EM fields can be described by the field strength (EB) and electromagnetic potentials (vector potential  $\mathbf{A}$  and scalar potential  $\phi$ ). However, only the former and the electromagnetic force when the carriers are passing through can be observed experimentally. Therefore, EB are regarded as the basic observables, and have physical meanings. The latter, electromagnetic potentials, are nothing but mathematical concepts that are introduced for convenient computation and do not have physical meanings. However, how can this explain the Aharonov-Bohm effect?

This effect mainly comes from the phase difference between the two paths induced by the local magnetic field. However, it seems that this phase difference is independent from the field strength in the paths, but the electromagnetic potentials that were thought of unphysical before. This, as we've already known, is equal the integral of a closed path. However, we also note that electromagnetic potentials changes along with the choice of gauge. Different gauge will give you different vector potential, then how could this work? Finally, we found that even though some vector potential will vanish depending on the choice of gauge, it cannot guarantee that the vector potential will vanish everywhere! We can still always find something that is independent from the local EM fields, but is gauge invariant on the electromagnetic potentials along the path. That's the integral of the closed path, which is gauge invariant.